Design Principles (July version)

Introduction

TS-EAS developed the following design principles in the first half of 2020 in order to reflect recent decisions and to guide future discussions for the development of EAD, EAC-CPF, and any other schemas that might be added to the EAS family in the future.

During the major revision of EAC-CPF (originally released in 2010, and revised slightly in 2018), it became apparent that the differences between EAD3 and EAC-CPF were difficult to keep track of not only for the maintainers of the standards but also for implementers. Early on, a major goal of the EAC-CPF revision was to align EAC-CPF more closely with EAD3. However, considering how different the two schemas actually are in their purpose and scope as well as in their maintenance setup, there was no guidance when it came to making decisions about how to harmonize the two. The following fifteen principles, focused on the schemas as well as the corresponding documentation, will help provide that guidance for TS-EAS going forward.

This document will be reviewed and updated, as necessary, based on feedback from the Committee (TS-EAS) and the Community, which includes archivists, software implementers, and anyone who interacts with the EAS schemas directly or the files associated with those schemas.

	Schema Design Principles		
	Principle	**Description**	**Examples**
1	Simplicity comes first.	The reason that simplicity is of such importance is because the Committee	EAD includes a filedesc element to encode bibliographic information about the finding aid itself, separate from the archival

		recognizes the high costs involved with implementing encoding standards. Therefore, our primary objective is to minimize the differences of the schemas that are managed by TS-EAS.	collection, whereas EAC-CPF does not provide this option. Adding filedesc to EAC-CPF would complicate EAC needlessly. Removing filedesc from EAD is an option, but that would only happen with a major revision of the EAD schema.
		Define as few elements and attributes as possible to lower costs related to implementation, training, and maintenance.	
		When choosing data types for elements and attributes, the schemas should favor the least restrictive but nevertheless most sensible option.	EAC-CPF has currently defined <recordId> as NMTOKEN, while EAD3 allows any text, enabling the use of blanks as well as additional special characters. In the context of reconciling both EAS, EAC-CPF will move towards the more open data type of EAD3.
		To ensure consistency, the schemas should be developed together rather than separately.	**Each element will be defined once, rather than multiple times in multiple places.**
		Use non-EAS namespaces sparingly.	We opted to remove the “xlink” attribute namespace from EAD3 and to define a subset of xlink-link attributes instead.
		But use external namespaces if another domain is to be included that is	LIDO’s use of the Geographic Markup Language, which imports the GML

		significantly outside the scope of archival description.	schema, rather than defining geographic elements in the LIDO namespace.
2	Community needs are tied with (and tied to) the first principle.	Although these principles are numbered, both the “first” principle and the “second” principle are of equal importance. In other words, Community needs (also) come first.	
		The Committee will receive and review feature requests from the Community on a continual basis, while also adhering to a predictable release schedule.	Feature requests can be submitted via GitHub directly or via the webform on the EAS microsite. Furthermore, TS-EAS has agreed an annual rolling revision cycle for clarity in timings..
		Each release will use “semantic versioning” numbers to clearly indicate to the Community what types of updates are included.	**The release of EAD3 v. 1.1.1 was a “patch” release that included backwards compatible bug fixes. For more information about semantic versioning scheme, see https://semver.org**
		Legacy practices of the Community and past decisions of the Committee should always be considered but rarely supersede the first principle.	
3	The schemas exist, first and foremost, to allow the Community to validate and share archival description, which should in turn adhere to archival descriptive standards.	The EAS schemas should be able to encode any description written according to archival standards.	The Committee focuses primarily on the descriptive standards produced by the International Council on Archives (ICA) and its expert groups.

		The schemas are intended to encode archival description and will not necessarily be able to encode other descriptive standards.	
		The schemas should not be made more complex to be able to encode non-archival description.	
		When necessary, the description should point to non-archival descriptions in other formats rather than embed them in EAS.	The <relations> section in the EAS would be the most prominent place to point to non-archival descriptions in other formats
		The schemas should support machine indexing as well as presentation of archival description.	
		The schemas should also permit the encoding of administrative metadata about the history of the archival description.	
4	Readability matters, especially since our Community is a community of people.	To help implementers as well as developers, the names of elements and attributes should be easy to read and to type.	
		Elements and attribute names are defined using a consistent naming scheme.	

		Elements and attributes are defined with English-language names.	
		The schemas will not reuse the same name for both an attribute and an element.	
5	Since our Community is an international community, the schemas support internationalization.	**Whenever repetition or internationalization is required, the committee will define those concepts as elements rather than attributes.**	
		The schemas must provide clear mechanisms to define the languages of what is being described (e.g. language of materials, language spoken, etc.) as well as the languages used in that description (e.g. the ability to encode multiple descriptive titles in different languages)	
		Recognizing language and cultural differences, the schemas will be flexible in how they permit the encoding of names, addresses, dates, time zones, typographical representations, language-based annotations, and more.	
		The Committee will review additional recommendations focused on supporting internationalization.	For example, the W3C's "Best Practices for XML Internationalization": https://www.w3.org/TR/xml-i18n-bp/

6	The schemas permit customizations, acknowledging that local requirements exist, without sacrificing interoperability.	**Allow minor expansions without invalidating the core schemas. This will reduce instances of "tag abuse," where implementers use tags and attributes in ways that contradict their definitions.**	Going forward, the schemas will permit attributes in any namespace other than the EAS namespace. When an implementer wants to use the complete set of XLink attributes, s/he could do so following this new approach, even with XLink having been removed from the core schemas.
		Additional XML validation options that cannot be easily defined or maintained in the core schemas will be supplied with Schematron.	EAD3 provides validation of controlled lists, such as ISO 639 language codes, in a separate Schematron file rather than in the core EAD3 schema.
		Encourage implementers to create their own local validation rules that restrict, rather than expand, the core schemas.	Making an optional element, such as "unittitle", a required element by using Schematron.
7	The schemas are free to use without restrictions on their use.	The schemas are published online.	
		The EAS schemas are currently serialized in RNG and XSD formats, both of which can be utilized with cost-free software.	
		The schemas are released under a CC0 license.	
8	The schemas will be revised, not in isolation, but in the context of related standards as well as changing technologies.	Additional EAS schema serializations may be explored by the Committee and provided if requested by the Community.	

		Updates to existing as well as emerging related standards may be explored by the Committee with regard to their relationship to and potential impact on EAS.	
		When defining new elements and attributes, the Committee will look at related standards for inspiration and may decide to adopt existing solutions rather than creating their own.	
	# Documentation Design Principles		
	Principle	**Description**	**Examples**
1	Documentation should be published online and made available under the terms of an open licence.	This includes the tag libraries, best practice guides, and any other usage documentation produced by EAS.	The EAD3 tag library is published under a Creative Commons CC BY licence.
2	Documentation should include encoding examples.	Examples should always be practical, providing useful and non-contradictory implementation guidance.	
		Examples should come from real-world implementations when possible.	

3	Documentation should be accessible, adhering to established standards and best practices.	PDF documentation should adhere to PDF/UA guidelines.	Generating PDF documentation using the `fop -a` option
		HTML documentation should adhere to WCAG AA guidelines.	
4	Documentation will be published in English as its primary language.	Suggested procedures for translating and generating translated documentation will be maintained and provided alongside source documents.	
5	Documentation should comprehensively reflect the most current version of the schemas.	Release documentation updates alongside schema updates.	
		All elements and attributes should be documented.	
6	Documentation should be transparent.	Explicitly list content changes made to documentation in appendices.	
		Use version control systems, such as Git, for maintaining documentation and publish changes in a public repository.	Publishing documentation on the Tag Libraries GitHub repository.
7	Documentation should be useful for both technical implementers and metadata creators.	Provide clear and specific narrative descriptions of how each element and attribute should be used in archival description.	
		Provide detailed technical expectations, requirements (repeatability, obligation,	

		data types, etc.), and uses for each element and attribute.	
		Provide references to other related SAA and ICA standards when possible.	